University of California, Santa Cruz
Department of Applied Mathematics and Statistics
Baskin School of Engineering
Classical and Bayesian Inference - AMS 132

Homework 1

Instructions: You have until Friday, January 26, to complete the assignment. It has to be returned during 10 last minutes of class (4:55 pm to 5:05 pm) or between 1:00 pm and 3:00 pm.

1. Suppose that the number of defects on a roll of magnetic recording tape has a Poisson distribution for which the mean θ is unknown.
   (a) Write the statistical model.
   (b) Assume that the mean number of defects on a roll of magnetic recording tape is either 1.0 or 1.5, and the prior p.f. of θ is $\xi(1.0) = 0.4$ and $\xi(1.5) = 0.6$. If a roll of tape selected at random is found to have three defects, what is the posterior p.f. of θ?
   (c) Under the same assumptions of (b), find the conditional p.f. of X_2 given $X_1 = 3$.

2. Suppose that a single observation X is to be taken from the uniform distribution on the interval $[\theta - 1/2, \theta + 1/2]$, where the value of θ is unknown, and the prior distribution of θ is the uniform distribution on the interval $[10, 20]$. If the observed value of X is 12, what is the posterior distribution of θ?

3. Automobile engines emit a number of undesirable pollutants when they burn gasoline. One class of polutants consists of the oxides of nitrogen. Suppose that $X_1, \dots, X_n$ describe the emissions of oxides of nitrogen, that can be described by a normal distribution with unknown mean θ and known variance σ^2. Also, assume that a normal prior p.d.f. for θ is considered with mean μ_0 and variance v_0^2.
   (a) Show that the posterior distribution of θ is a normal distribution with mean and variance given by $\mu_1 = \frac{\sigma^2\mu_0 + n v_0^2 \overline{x}_n}{\sigma^2 + n v_0^2}$, and $v_1^2 = \frac{\sigma^2 v_0^2}{\sigma^2 + n v_0^2}$.
   (b) Find the distribution of X_{n+1} given $X_1 = x_1, \dots, X_n = x_n$.
   (c) Assume that $\sigma = 0.5$, $\mu_0 = 2.0$, and $v_0 = 1.0$. In a random sample of $n = 46$ engines, the average of the emissions of oxides of nitrogen was 1.329. Compute $P(X_{47} < 1.5 \mid x_1, \dots, x_n)$.

4. Suppose that $X_1, \dots, X_n$ form a random sample from a distribution for which the p.d.f. $f(x \mid \theta)$ is $f(x \mid \theta) = \theta e^{-\theta x}$, for $x > 0$, and $f(x \mid \theta) = 0$, otherwise. Suppose that the value of the parameter θ is unknown ($\theta > 0$), and the prior distribution of θ is the gamma distribution with parameters α and β ($\alpha > 0$ and $\beta > 0$).
   If $Y \mid a, b \sim Gamma(a, b)$, then $f(y \mid a, b) = \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by}$, $E(Y \mid a, b) = a/b$ and $Var(Y \mid a, b) = a/b^2$.

(a) Write the statistical model.

(b) Determine the mean and the variance of θ after observing $X_1 = x_1, \ldots, X_n = x_n$.

(c) If $\alpha = \beta = 0$, is the prior p.d.f. of θ improper? Is the posterior p.d.f. of θ improper?

5. Show that the family of beta distributions with parameters $a, b > 0$, is a conjugate family of prior distributions for samples from a negative binomial distribution with a known value of the parameter r and an unknown value of the parameter p $(0 < p < 1)$.
If $X \mid p \sim NegativeBinomial(r, p)$, then $P(X = x \mid p) \propto p^r(1-p)^x$.

6. Suppose that the time in minutes required to serve a customer at a certain facility has an exponential distribution for which the value of the parameter θ is unknown and that the prior distribution of θ is a gamma distribution.
If $Y \mid a, b \sim Gamma(a, b)$, then $f(y \mid a, b) = \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by}$, $E(Y \mid a, b) = a/b$ and $Var(Y \mid a, b) = a/b^2$.

   (a) The average time required to serve a random sample of 20 customers is observed to be 3.8 minutes. If the prior distribution of θ has mean equal to 0.2 and the standard deviation equal to 1, what is the posterior distribution of θ?

   (b) For a distribution with mean $\mu \neq 0$ and standard deviation $\sigma > 0$, the *coefficient of variation* of the distribution is defined as $\sigma / \mid \mu \mid$. Suppose that the coefficient of variation of the prior gamma distribution of θ is 2. What is the smallest number of customers that must be observed in order to reduce the coefficient of variation of the posterior distribution to 0.1 or less?

   (c) Suppose now that the coefficient of variation of the prior gamma distribution of θ is 0.2. What is the smallest number of customers that must be observed in order to reduce the coefficient of variation of the posterior distribution to 0.1?